Iraqi Journal of Biotechnology, 2016, Vol. 15, No. 2 , 1-8

Allele Frequency of 15 Autosomal Short Tandem Repeat (STR) Loci in Al Anbar - Iraqi Population

Thooalnoon Younes AL-Janabi, Mohammed Mahdi AL-Zubaidi, Dhuha Salim Namaa, Sura Nabeel Hameed, Asia Abd- Alatief Mahdi

Forensic DNA Center Research and Training, AL-Nahrain University

Received: November 4, 2015 / **Accepted**: May 18, 2016

Abstract: Allele frequencies for 15 STR loci (CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317,D16S539, D18S51, D21S11, FGA, TH01, TPOX, VWA, D2S1338, D19S433) included in the AmpFlSTR Identifiler kit were determined in a sample of 132 unrelated people originating From Al Anbar - Iraqi Population, samples were extracted using a Prep Filer Forensic DNA Extraction Kit.
DNA concentration measured using Nano drop Simultaneous amplifications of 15 STR loci and a gender determination marker (multiplexed PCR) were done by using the AmpFlSTR® Identifiler® PCR Amplification Kit according to the user's manual recommendations .
The separation and detection of amplified products were conducted with the 3130 xl Genetic Analyzer 16-capillary array system following manufacturer's protocols, A different number of alleles were observed with frequencies ranging between 0.0038 (FGA- allele 18, 19.2, 23.2, 27 and 28 - , D18S51 - allele 20 and 23, D19S433 - allele 9.2,11 and 12.2, D2S1338 – allele 27, D13S317 – allele 16, D3S1358 – allele 13 and 19 CSF1PO – allele 14, D21S11– allele 30.2 and 32 and D8S1179 - allele 9 and 17) and 0.5795 (TPOX-allele 8). The highest heterozygosity is observed for FGA (85.60 %) where the smallest heterozygosity value is obtained for TPOX (61.36 %). The loci were observed to have high discriminating power, as the power of discrimination of each loci varied from 0.812 (TPOX) to 0.973 (D2S1338). All loci but D21S11 (<0.001), D19S433 (<0.001), D5S818 0.00204 and FGA (<0.00024) met Hardy-Weinberg expectations (P > 0.05). Significant departure from Hardy Weinberg Equilibrium (HWE) expectations were observed in loci D21S11 (0.001), D19S433 (0.001), D5S818 (0.002) and FGA (0.00).
The results of the population genetics tests and pairwise comparisons suggest that these allele frequency databases are suitable for the purpose of identification in paternity or forensic investigations.

Key Words: DNA, STR, AL-Aanbar - Iraqi Population, AmpFlSTR pro®ler, Forensic identi®cation.

Corresponding author: should be addressed (Email: fdna_thro15@nahrainuniv.edu.iq)

Introduction

A shart tandem repeat (STR) in DNA occurs when a pattern of two or more nucleotides are repeated and the repeated sequences are directly next to each other. The pattern may range in length from 2 to 16 base pairs (bp) (such as (CATG) in a genomic region) and is typically in the non-coding introns region (1).
Introducing of a set of short tandem repeat (STR) loci as the markers induced a significant progress in this field of science (2,3,4). STR loci show variability among individuals in population and that makes these sequences important in genetic

mapping, linkage analysis, and identity testing in forensic cases, paternity testing, missing person's investigations, and mass disaster victim identification. In order to determine the probability of a particular genotype, population data must be gathered with a proper sample size to make an estimate of the frequency of each possible allele and genotype. Short tandem repeat (STR) loci are highly polymorphic and found in abundance throughout the human genome (5). These genetic markers have proven to be particularly well suited for medical research, ethnographic studies and for discrimination/individualization in a variety of forensic and judicial settings (6). The discriminating power of multiplexed STR markers is very high compared to RFLP methods, and with proper use of the population databases, estimates of match Probability approach 1 in one billion. Additionally typing of multiple loci can be accomplished in a single multiplex reaction (7). In STR analysis, a PCR is performed using primers on each side of the microsatellite, followed by electrophoresis and detection of fragment lengths. The STRs analyzed in forensic investigations are generally 150-400 bp long regions of tetra nucleotide repeats (8) Therefore, importance of understanding the used marker heterogeneity within different populations is constantly emphasized (9). Different number and different sets of STR loci in a different number of the individuals of southern origin were used in previous studies of Anbar - Iraqi Population.

In this study, allele frequencies for the fifteen autosomal STR loci namely D8S1179, D21S11, D7S820, CSF1P0, D3S1358, TH01, D13S317, D16S539, D2S1338, D19S433, vWA, TPOX, D18S51, D5S818, and FGA and Amelogenin.—were analyzed in 132 unrelated Anbar individuals.

The aim of this work was to establish a database of the Anbar- Iraqi Population for forensic purposes including paternity testing and to Determination The frequency of alleles at the locus beside to Determination the frequency of genotypes at the locus.

Population studies are not enough to make reliable interpretation about polymorphism of STR loci in Iraq. For this reason, it is necessary to do population studies on a province and region level and collect all data to form Iraqi database. In this study we present the allele frequencies and forensic efficiency values for the 15 loci in a sample of 132 unrelated Anbar- Iraqi Population.

Materials and Methods

Samples Collection

Buccal swabs (Sterile Omni Swab or Sterile Foam Tipped Swabs, Whatman International Ltd., Maidstone, UK) (10) were collected from 132 unrelated people originating in Al Anbar - Iraqi Population (Ramadi and Fallujah) on September 2013 from 82 men's and 50 women's all participants gave their information's after accepting to be subjects of this project information include : personal factor (gender , age , occupation, address ethnicity, Place of birth and if they are smokers, drinkers and family history of genetic diseases.

DNA Extraction

DNA samples were extracted using a Prepfiler Forensic DNA Extraction Kit (Applied Biosystems, Foster City, CA).

DNA concentration was measured with Nanodrop (Thomson, Wilmington, DE).

PCR

Simultaneous amplifications of 15 STR loci and a gender determination marker (multiplexed PCR) were done by using the AmpF*l*STR® Identifiler® PCR Amplification Kit according to the user's manual recommendations (AmpFlSTR1 Identifiler TM PCR Amplification Kit User's Manual, 2001) (11).
The 15 loci amplified in this studies are D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317, D7S820, TH01, TPOX, CSF1PO, D19S433, D2S1338, D16S539 and Amelogenin (gender determination marker), The samples were amplified with an Applied Biosystems Veriti® PCR System and process positive and negative control to get approve if the products is acceptable or not and if there any problem in samples or kit.

Typing

The separation and detection of amplified products were conducted with the 3130 XL Genetic Analyzer 16-capillary array system (Applied Biosystems, Foster City, CA, USA) following manufacturers
Protocols, Amplification products were diluted 1:15 in Hi- Di™ formamide and GS500-LIZ internal size standard (Applied Biosystems) POP™-4 (Applied Biosystems) was used for higher resolution separations through a 36 cm array.
Data collections were done using Data Collection v3.0 software (Applied Biosystems, Foster City, CA, USA) and samples were analyzed by Gene Mapper v. 1.1 software (Applied Biosystems, Foster City, CA, USA) the homozygous min peak height is 250 and the heterozygous min peak height is 150, max peak height is 5000.

Statistical Analyses

Several forensic and population parameters such as the power of discrimination (PD), the a priori chance of exclusion (CE), the polymorphism information content (PIC), the paternity index (PI), the random match probability (RMP) and marker's observed and expected heterozygosity (Ho and He, respectively) of the 15 loci
Were estimated and P value of exact test for Hardy–Weinberg equilibrium and expected heterozygosity by using the manual method.

Results and Discussion

Anbar is a province of Iraq and is located in the west of the country. Is the largest of Iraq's provinces join the space where the equivalent of one- third of the area of Iraq. An area of 138.5 thousand square kilometers, and has a total population of 1900000 people (almost two million) (July 2013) From Wikipedia .Before bordered on the north provinces of Salahuddin and Nineveh, the Syrian Arab Republic from the north-west , Jordan from the West, The province of Baghdad from the east South of Saudi Arabia and from the south-east provinces of Karbala and Najaf.
The observed allele frequencies for the 15 STR loci and results of forensic efficiency limits for Al Anbar - Iraqi Population are shown in Tables 1 and 2. A different number of alleles were observed with frequencies ranging

between 0.0038 (FGA- allele 18, 19.2,23.2, 27 and 28 - , D18S51 - allele 20 and 23, D19S433 - allele 9.2,11 and 12.2, D2S1338 – allele 27, D13S317 – allele 16, D3S1358 – allele 13 and 19 CSF1PO – allele 14, D21S11– allele 30.2 and 32 and D8S1179 - allele 9 and 17) and 0.5795 (TPOX-allele 8). The highest heterozygosity is observed for FGA (85.60 %) where the smallest heterozygosity value is obtained for TPOX (61.36 %). The loci were observed to have high discriminating power, as the power of discrimination of each loci varied from 0.812 (TPOX) to 0.973 (D2S1338). All loci but D21S11 (<0.001), D19S433 (<0.001), D5S818 0.00204 and FGA (<0.00024) met Hardy- Weinberg expectations (P > 0.05) (12).

Where the PIC ranged from 0.57 (TPOX) to 0.86 (D2S1338). The combined power of discrimination for the 15 STR loci studied is 0.999 in Table 3 that should be explain for the identification of any person even for an extremely large population size. All 15 loci give a combined chance of exclusion in no paternity of 99.9%. The Combined Exclusion Probability for the 15 STR loci studied is (0.99).

In the end we noticed an 11 locus are non-significant when compared with Hardy- Weinberg (P > 0.05) and the locus are D8S1179, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, D2S1338, vWA, TPOX and D18S51.

This difference represents Deviations from the Hardy–Weinberg Equilibrium (p < 0.05) for this study when comparing with the other population studies, and this Deviations Represent a variation for this region.

Table 1. Observed allele frequency distributions of the 15 STR loci in 132 unrelated Anbar - Iraqi Population

D8S1179

Allele	Freq *
8	0.0227
9	0.0038
10	0.0985
11	0.1023
12	0.0833
13	0.2500
14	0.1742
15	0.2159
16	0.0455
17	0.0038

**** P: 0.830**

D21S11

Allele	Freq
26	0.0114
27	0.0114
28	0.1439
29	0.2652
29.2	0.0189
30	0.2197
30.2	0.0038
31	0.0606
31.2	0.1023
32	0.0038
32.2	0.1364
33	0.0000
33.2	0.0152
34.2	0.0076

P <0.001

D7S820

Allele	Freq
7	0.0076
8	0.1515
9	0.0985
10	0.3258
11	0.2538
12	0.1288
13	0.0227
14	0.0114

P 0.892

CSF1PO

Allele	Freq
9	0.0114
10	0.2879
11	0.2879
12	0.3902
13	0.0189
14	0.0038

P 0.857

FGA

Allele	Freq
18	0.004
19	0.098
19.2	0.004
20	0.087
21	0.174
21.2	0.064
22	0.087
23	0.140
23.2	0.004
24	0.152
25	0.140
26	0.030
27	0.004
28	0.004
29	0.008

P 0.000

D3S1358

Allele	Freq
13	0.0038
14	0.0265
15	0.2841
16	0.3333
17	0.2273
18	0.1212
19	0.0038

P 0.975

TH01

Allele	Freq
6	0.3371
7	0.2008
8	0.1136
9	0.1780
9.3	0.1553
10	0.0152
11	0.0000

P 0.352

D13S317

Allele	Freq
8	0.1553
9	0.0720
10	0.0455
11	0.3371
12	0.2992
13	0.0492
14	0.0379
16	0.0038

P 0.378

D16S539

Allele	Freq
8	0.0189
9	0.1932
10	0.0758
11	0.2765
12	0.2008
13	0.2197
14	0.0152

P 0.157

D2S1338

Allele	Freq
16	0.0341
17	0.1742
18	0.1402
19	0.1136
20	0.1667
21	0.0833
22	0.0455
23	0.0947
24	0.0758
25	0.0606
26	0.0076
27	0.0038

P 0.094

D19S433

Allele	Freq
9.2	0.0038
11	0.0038
12	0.1023
12.2	0.0038
13	0.2348
13.2	0.0227
14	0.2462
14.2	0.0606
15	0.1250
15.2	0.0871
16	0.0379
16.2	0.0568
17.2	0.0152

P <0.001

vWA

Allele	Freq
14	0.0417
15	0.0455
16	0.2652
17	0.3068
18	0.2311
19	0.1023
20	0.0076

P 0.656

TPOX

Allele	Freq
8	0.5795
9	0.1364
10	0.0871
11	0.1553
12	0.0417

P 0.518

D18S51

Allele	Freq
10	0.023
11	0.019
12	0.129
13	0.197
14	0.189
15	0.098
16	0.102
17	0.106
18	0.072
19	0.027
20	0.004
21	0.023
22	0.008
23	0.004

P 0.058

D5S818

Allele	Freq
8	**0.0076**
9	0.0492
10	0.0833
11	0.3485
12	0.3144
13	0.1780
14	0.0189
15	0.0000

P 0.002

***Allele frequency.**
**** P.value.**

Table 2:Observed and expected Homozygosity and heterozygosity with power of discrimination of the 15 STR loci in 132 unrelated Anbar - Iraqi Population

Locus	Observed Homozygosity	Observed2 Heterozygosity	Expected Homozygosity	Expected3 Heterozygosity	Power of Discrimination
D8S1179	19.70%	80.30%	16.92%	83.08%	0.950
D21S11	21.21%	78.79%	17.30%	82.70%	0.948
D7S820	23.48%	76.52%	22.05%	77.95%	0.919
CSF1PO	34.85%	65.15%	31.85%	68.15%	0.834
D3S1358	27.27%	72.73%	25.89%	74.11%	0.888
TH01	21.21%	78.79%	22.29%	77.71%	0.917
D13S317	26.52%	73.48%	23.84%	76.16%	0.908
D16S539	23.48%	76.52%	20.87%	79.13%	0.924
D2S1338	15.15%	84.85%	11.93%	88.07%	0.974
D19S433	18.18%	81.82%	15.86%	84.14%	0.957
vWA	31.82%	68.18%	23.22%	76.78%	0.909
TPOX	38.64%	61.36%	38.79%	61.21%	0.813
D18S51	17.42%	82.58%	13.00%	87.00%	0.970
D5S818	34.09%	65.91%	26.18%	73.82%	0.889
FGA	14.39%	85.61%	12.27%	87.73%	0.972

Table 3: Forensic efficiency parameters for the 15 STR loci in 132 unrelated Anbar - Iraqi Population *

Statistical parameter

Locus	Matching Probability	Polymorphism Information Content	Power of Exclusion	Typical Paternity Index	Paternity Index Expected	Power of Discrimination
D8S1179	0.050	0.809	0.563	2.538	2.955	0.950
D21S11	0.052	0.805	0.539	2.357	2.891	0.948
D7S820	0.081	0.747	0.505	2.129	2.268	0.919
CSF1PO	0.166	0.617	0.362	1.435	1.570	0.834
D3S1358	0.112	0.696	0.453	1.833	1.931	0.888
TH01	0.083	0.744	0.539	2.357	2.243	0.917
D13S317	0.092	0.726	0.463	1.886	2.097	0.908
D16S539	0.076	0.759	0.505	2.129	2.396	0.924
D2S1338	0.026	0.869	0.641	3.300	4.190	0.974
D19S433	0.043	0.823	0.588	2.750	3.153	0.957
vWA	0.091	0.731	0.396	1.571	2.154	0.909
TPOX	0.187	0.575	0.322	1.294	1.289	0.813
D18S51	0.030	0.856	0.601	2.870	3.846	0.970
D5S818	0.111	0.695	0.370	1.467	1.910	0.889
FGA	0.028	0.865	0.656	3.474	4.076	0.972
Combined Matching Probability (CMP)		4.05415E-18				
Combined Discrimination Power (CDP)		0.99999				
Combined Exclusion Probability (CEP)		0.986719584				

References

1- John, B. (2005). Forensic DNA typing second edition.
2- Gill, P; Kimpton C.P; Urquhart, A.; Oldroyd, N.; Millican, E.S.; Watson, S.K. and Downes, T.J. (1995). Automated short tandem repeat (STR) analysis in forensic casework – a strategy for the future. , 16: 1543-52.
3- Thomson, J.A; Pilotti, V.; Stevens, P.; Ayres, K.L; and Debenham, P.G. (1999). Validation of short tandem repeat analysis for the investigation of cases of disputed paternity. Forensic Sci Int., 100: 1-16.
4- Alonso, A.; Anđelinović, S.; Martin, P.; Sutlović, D.; Erceg, I. and Huffine, E. (2002). Croat Med J. DNA typing from skeletal remains: evaluation of multiplex and megaplex STR systems on DNA isolated from bone and teeth samples.,42: 260-266
5- Yasin, S.R.; Hamad, M.M.; ElKarmi, A.Z. and Jaran, A.S. (2005). Croatian Medical Journal. African Jordanian population genetic database on fifteen short tandem repeat genetic loci.
6- Tuğ, A.; Erkol, Z.; Çetinyürek, A.; Alakoç, Y.D.; Elma, C. and Büken, B. (2010). Turkish J Med Sci. Allele distribution data for 16 short tandem repeat loci in Bolu. The Scientific and Technological Research Council of Turkey. , 40 (4):659–64.
7- Hochmeister, M.N.; Budowle, B.; Jung, J.; Borer, U.V.; Comey, C.T. and Dirnhofer, R. (1991). International Journal of Legal Medicine. PCR-based typing of DNA extracted from cigarette butts. , 104(4): 229-33.
8- Gill, P. (2002). Bio Techniques. Role of short tandem repeat DNA in forensic casework in the UK –past, present, and future perspectives. , 32(2): 366-385.
9- Cohen, J.E. (1990). Am. J. Hum. Genet. DNA fingerprinting for forensic identification: potential effects on data interpretation of subpopulation heterogeneity and band number variability. , 46: 358-68.
10- Sterile Omni Swab or Sterile Foam Tipped Swabs, Whatman International Ltd., Maidstone, UK.
11- Applied Biosystems. AmpFlSTR Identifiler PCR Amplification Kit User's Manual, Foster City, CA. 2001; 4323291.
12- Lewis, P.; Zaykin, D. (2001). (Internet). Genetic Data Analysis: Computer program for the analysis of allelic data.